AMENDMENTS TO THE CLAIMS

This listing of claims will replace all prior versions of claims in the application:

Listing of Claims:

1. (Currently Amended) A system that facilitates extracting data in connection with spam processing, comprising:

a computer readable storage medium comprising:

a component that receives a message and extracts a set of features associated with some part, content or content type of a message; and

an analysis component that ~~at least~~ examines consecutiveness of characters within a subject line of the message ~~or at least examines~~ and a content type of the message for spam in connection with building a filter, wherein the content type is case-sensitive, comprises primary content-type and a secondary-content type, or combinations thereof.

2. (Original) The system of claim 1, the analysis component determines frequency of consecutive repeating characters within the subject line of the message.

3. (Original) The system of claim 2, the characters comprise letters, numbers, or punctuation.

4. (Original) The system of claim 1, the analysis component determines frequency of white space characters within the subject line of the message.

5. (Original) The system of claim 1, the analysis component determines distance between at least one alpha-numeric character and a blob.

6. (Original) The system of claim 1, the analysis component determines a maximum number of consecutive, repeating characters and stores this information.

7. (Original) The system of claim 1, the analysis component establishes ranges of consecutive, repeating characters, the ranges corresponding to varying degrees of spaminess, whereby messages can be sorted by their respective individual count of consecutive repeating characters.

8. (Cancelled)

9. (Previously Presented) The system of claim 1, the analysis component compares the content type of a current message to stored content types of a plurality of other messages to facilitate determining whether the message is spam.

10. (Cancelled)

11. (Cancelled)

12. (Original) The system of claim 1, the analysis component further determines time stamps associated with the message.

13. (Currently Amended) The system of claim 12, the analysis component ~~determining~~determines a delta between time stamps.

14. (Original) The system of claim 13, the delta is between a first and a last time stamp.

15. (Original) The system of claim 1, the analysis component determines at least one of: a percentage of white space to non-white space in the subject line of the message and a percentage of non-white space and non-numeric characters that are not letters in the subject line of the message.

16. (Original) The system of claim 1, the filter being a spam filter.

17. (Original) The system of claim 1, the filter being a parental control filter.

18. (Original) The system of claim 1, further comprising a machine learning system component that employs at least a subset of extracted features to learn at least one of spam and non-spam.

19. – 41. (Cancelled)

42. (New) A method for evaluating spam as a function of message content, comprising:

employing a processor executing computer readable instructions stored on a computer readable storage medium to implement the following:

parsing a message to extract a set of features associated with a part, content, or content type of the message, wherein the content type is case-sensitive and comprises a primary content-type and a secondary-content type, or combinations thereof; and

examining the extracted set of features and consecutiveness of repeating characters within a subject line of the message to classify the message as spam or not spam; and

processing the message as a function of the classification.

43. (New) The method of claim 42, examining the consecutiveness of repeating characters comprises determining a frequency of the consecutive of repeating characters, wherein the characters comprise letters, numbers, punctuation, or white space.

44. (New) The method of claim 42, examining the extracted set of features comprises determining a distance between at least one alpha-numeric character and a blob.

45. (New) The method of claim 42, further comprising:

establishing ranges of consecutive, repeating characters, the ranges correspond to various degrees of spaminess; and

employing the ranges to sort messages by their respective individual count of consecutive repeating characters.

46. (New) The method of claim 42, further comprising comparing the set of features of the message to stored content types of a plurality of other message to determine whether the message is spam.